REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A240 740

...ted to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and ...lection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including ... Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302 ... on Project (0704-0188), Washington, DC 20503.

	2 REPORT DATE August 1991	3 REPORT TYPE AND DATES COVERED professional paper
4 TITLE AND SUBTITLE MODELING NATURAL LANGUAGE FOR USE IN THE COMBINATION OF EVIDENCE PROBLEM		5 FUNDING NUMBERS In-house funding
6 AUTHOR(S) I. R. Goodman		
7 PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Ocean Systems Center San Diego, CA 92152-5000		8 PERFORMING ORGANIZATION REPORT NUMBER
9 SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Ocean Systems Center San Diego, CA 92152-5000		10 SPONSORING/MONITORING AGENCY REPORT NUMBER
11 SUPPLEMENTARY NOTES		
12a DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		12b DISTRIBUTION CODE

13 ABSTRACT (Maximum 200 words)

The general problem of combining information from various sources, linguistic and/or numerical, has been treated in previous work. This paper continues the effort with emphasis on the modeling of linguistic information. Included in the results is a new approach to the modeling of conditional expressions. In addition, temporal and modal operators are treated as well as linguistic connectives and relations.

Published in the *Proceedings of the 7th MIT/ONR Workshop on C³ Systems*, December 1984.

14 SUBJECT TERMS linguistic connectives, algorithm, probabilistic			15 NUMBER OF PAGES
			16 PRICE CODE
17 SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18 SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19 SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20 LIMITATION OF ABSTRACT SAME AS REPORT

NSN 7540-01-280-5500

Standard form 298

UNCLASSIFIED

21a NAME OF RESPONSIBLE INDIVIDUAL	21b TELEPHONE (Include Area Code)	21c OFFICE SYMBOL
I. R. Goodman	(619) 553-4014	Code 421

NSN 7540-01-280-5500

Standard form 298

UNCLASSIFIED

DECEMBER 1984 LIDS-R-1437

PROCEEDINGS OF THE

7TH MIT/ONR WORKSHOP ON C^3 SYSTEMS

EDITED BY:
MICHAEL ATHANS
ALEXANDER H. LEVIS

Laboratory for Information and Decisions Systems
MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 021390

MODELING NATURAL LANGUAGE INFORMATION FOR USE IN THE COMBINATION OF EVIDENCE PROBLEM

I.R. Goodman

Command and Control Department, Code 421,
Naval Ocean Systems Center, San Diego, California 92152.

Abstract

The general problem of combining information from various sources, linguistic and/or numerical, has been treated in previous work. This paper continues the effort with emphasis on the modeling of linguistic information. Included in the results is a new approach to the modeling of conditional expressions. In addition, temporal and modal operators are treated as well as linguistic connectives and relations.

INTRODUCTION

This paper continues the ongoing investigations into the general problem of combining information that arises from many different sources in location and type, the latter including a wide variety of linguistic-based or probabilistic origins. A major application of this work has been to the target tracking and data association- or "correlation"- problem in Naval Ocean Surveillance. (For previous studies see, e.g., [1],[2],[3].)

In brief, combination of evidence has been considered from a multiple-valued logical viewpoint, related to the theory of fuzzy sets as originally developed and refined by Zadeh [4],[5] and carried further by other researchers[6],[7],[27]. Part of the work has also been motivated and guided by parallel results from the theory of random sets and "flou" sets [8],[9]. The major application to the problem of data association has culminated in the PACT (Possibilistic Approach to Correlation and Tracking) algorithm [2],[3]. This is a procedure which operates upon raw data or suitably smoothed or predicted data, depending upon the given scenario, which is categorized according to attribute type. Thus for example, data may be classified into geolocation- the most common and first historically to be considered by researchers in correlation theory (see the extensive work involving only geolocation attributes in the Naval Ocean Surveillance Correlation Handbook [10],[11])-as well as other categories. The latter include classification of various sorts, visual sighting information, and many kinds of sensor system state parameter data. In addition, PACT operates upon predetermined functions or tables of possible errors involved in the reporting of the attribute information, as well as upon a collection of inference rules obtained from a combination of experts and analytic considerations. These inference rules connect degrees of possible matches of information for various combinations of attributes between a typical pair of track histories considered for correlation and the resulting levels of possible correlations. Thus, a typical inference rule might state in its linguistic form "If geolocation information matches to degree 0.8 and radar (of type A, say) parameter information matches rather poorly, then the correlation level between the two track histories is low, but not impossible." Some of the attributes in the above process are statistical in nature, resulting in statistical procedures for the modeling of associated error tables and matching tables used in the inference rules, while others are more subjective in nature and are obtained from pooling panels of experts or from other non-analytic sources. Examples of the former include geolocation and sensor state attributes, while examples of the latter include various verbal descriptions, classifications, and intelligence information. The output of the PACT algorithm is a posterior possibility or membership function of the correlation level between any two given track histories of interest. In turn, a single figure-of-merit has been developed which represents the overall correlation level between the two histories [12]. Also, a scheme may be established in which the algorithm acts as a "black box" producing a table of correlation values relative to all tracks of interest to be used in performing correlation and tracking. Two major breakthroughs have been recently obtained concerning the design and asymptotic behavior of the algorithm. It has been shown that guidelines may be established for the choice of class of operators used in the algorithm and that if relatively data matches occur, the algorithm will yield asymptotically consistent correlation levels. Conversely, low data matches drive the algorithm eventually to low correlation outputs, as should be intuitively.(See [12] for both results.)

Since the core of the approach to combining evidence, as outlined above, consists of inference rules and error tables modeled within a framework of multiple-valued logic, modeling of these relations is critical. In addition, it is important that a wide scope of relations can be treated. However, at present, the PACT algorithm can only handle the simplest type of inference rules. The following inference rule is far too complex in form to be incorporated in the algorithm: S^*= "In Ocean region V, and usually in Region W, if the weather is poor and the sea state corresponds to relatively high turbulance, then indications by sensor system A that a submarine was in the area are not that reliable and probably should be discounted in favor of geolocation matching information obtained from sensor system B, although exceptions to this can occur when visibility is up to about two miles, in which case it has been shown that A-data matching should be assigned a much higher degree of importance in its effect upon correlation." Another example is given by: S^{**}="Contact with the ship was held for about two hours, but was lost just before the Straights of Skagerrak were sighted, although purple side-insignias may have been spotted as well as an oval-shaped dome near the rear of the ship, but a foggy condition pervaded the area preventing any further identification."

Such examples as above illustrate the typical problems faced in modeling and symbolizing natural language. This includes the interpretation of modifiers such as "usually", "relatively high", "not that reliable", modal and temporal operators such as "probably should be","was held for about two hours",and verb/predicative relations such as "Contact with the ship", "was lost", "foggy condition pervaded the area". (See [13],[14] for related linguistic problems arising in expert systems.)

A systematic approach to the full symbolization of language is thus most desirable. In this paper, some modest efforts in this direction are made. Conditional expressions, such as "most tall ships in region 5 are enemy ones of type F" are considered. The approach here is in contrast to Zadeh's rather arbitrary "fuzzy cardinality" approach which cannot be directly derived from multivalued truth considerations [5],[15]. A comprehensive approach is taken to the modeling of temporal and modal relations, extending earlier ideas of Zadeh's PRUF technique [5]. Some examples are presented illustrating these ideas with an important application to the combination of evidence procedure as given in PACT. [16] contains a number of other modeling procedures in addition to those presented in this paper.

NATURAL LANGUAGE, FORMAL LANGUAGE, AND SEMANTIC EVALUATIONS

Too often in the past, natural language information was neglected in favor of "more precise" numerical data. Or, such information at times was arbitrarily precisified to be in numerical form. Since the onset of Chomsky and others, more rigorous outlooks have been taken toward the understanding and modeling of information content in language [17],[18]. With the work of Zadeh on PRUF [5] began a new era in the development of a calculus for semantic evaluation of natural language. This section continues in the direction of Zadeh. The following basic premises are assumed:

(a) All natural language information is translatable into sequences of English sentences. The problem of whether a given natural language molds the speaker's thoughts due to its structure and limitations-the Whorf-Sapir hypothesis, or whether this is not valid-as Berlin and Kay claim (see [19] for comments)will not be dealt with here.

(b)Ambiguity of meaning is expressed by (subjectively) weighting the possibility of interpretations. Thus, e.g. the expression "I like her well." could be

S_1="I really like her."

S_2="I wish she remains well (in good health).

S_3="I want her to become well."

S_4="I like the well that she owns."

Weight w_i could be assigned to S_i , $i=1,\ldots,4$. Usually context allows for resolution of these possible branches of meaning. For simplicity, it will be assumed here no ambiguity is present. (See [19] for further discussion.)

(c)Any given sentence in actuality represents an equivalence class of possibly differently appearing-i.e., syntactically different- sentences, all having the same semantic evaluation, a number lying in the unit interval [0,1] representing its truth value. This is related to Chomsky's concepts of transformational generative grammar, where changes in forms of sentences are due to word order rearranged, use of synonyms, change of voice from active to passive, or other superficialities[17],[18]

(d)Parsing Principle: Given any sentence (or any equivalence class of sentences) there exists an analytic form or parsing which is semantically the same but is structured within a formal language. This is related to Chomsky's deep structure analysis [17],[18]. (For further details on formal language and multivalued logic,see [21],[27],[16]).Attempts at establishing automatic procedures for parsing natural language into a corresponding formal language form are many and the area remains a lively one for research. (See the large compendium of approaches in [20].)

A typical parsing analysis yields for any given compound sentence S

$$S=\mathrm{comb}(\ldots,\mathrm{not},\&,\mathrm{or},\mathrm{if}()\mathrm{then},\ldots)(S_1,S_2,\ldots,S_r) \quad (1)$$

where the operators "not","&","or", etc. all indicate the usual unary or binary linguistic connectors and comb indicates some sequential combination of these connectors with sentences $S_1,S_2,\ldots,S_r$, the latter all having simpler forms than S does. In turn, each S_i also has a parsed form in terms of relatively simpler sentences, etc.

(e) Modified Principle of Abstraction: Any sufficiently simple sentence,such as the components S_i in (1), has a unique corresponding semantically equivalent form

$$(x \in A) \quad (2)$$

where A represents a generalized set, property, or attribute, x is a possible vector of elements in the ordinary sense, and ϵ is the extended set membership relation for generalized sets. (See [16] for further details on these relations.) As in ordinary set relations, A is considered a subset of an ordinary set X called the universe of discourse or base space and in the ordinary sense,x is in X, i.e. $x \in X$. It should be noted that this apparently reasonable principle can lead to paradoxes in formal logical systems , such as in classical naive set theory or even in set theory based on multiple-valued logic, for a wide variety of logics (except for Lukasiewicz-$\aleph_1$Logic- see [27]). In the work here, these diffficulties will be ignored for the time being.

Thus, (1) and (2) yield for sentence S

$$S=\mathrm{comb}(\ldots,\mathrm{not},\&,\mathrm{or},\ldots)(x_1\epsilon A_1,x_2\epsilon A_2,\ldots,x_r\epsilon A_r)$$
$$=(x \in A) \quad (3)$$

where

$$x=(x_1,x_2,\ldots,x_r) \ , \quad (4)$$

$$A=\mathrm{comb}'(\ldots,\mathsf{C},\mathsf{X},\dagger,\nrightarrow,\ldots)(A_1,A_2,\ldots,A_r), \quad (5)$$

where comb' is some other combination function and C is the complement operator on generalized sets, corresponding to "not" ,
X is the cartesian product operator corresponding to "&",
$\dagger$ is the cartesian sum operator corresponding to "or", etc. (Again, see [16] for further details.)

(f) Principle of Semantic Evaluation: Any sentence S has a truth value $||S||$, a number in [0,1] which can be evaluated through the values of the semantic function $||\cdot||$ over component parts of S,given the particular semantic function, or equivalently , logic chosen [21].

(i) If the semantic function is truth functional, then eq.(3) is evaluated as

$$||S||=\mathrm{comb}(\ldots,\phi_N,\phi_\&,\phi_0,\phi_\nrightarrow,\ldots)(\phi_{A_1}(x_1),\ldots,\phi_{A_r}(x_r))$$
$$=\phi_A(x) \ , \quad (6)$$

where

$$\phi_{A_i}(x_i) = ||x_i \in A_i||, i=1,\ldots,r, \quad (7)$$

yielding in general the membership or possibility function $\phi_{A_i}:X_i \to [0,1]$, and where

$$\phi_N = ||\mathrm{not}||:[0,1] \to [0,1] \text{ is a nonincreasing} \quad (8)$$

function with $\phi_N(0)=1$ and $\phi_N(1)=0$, the classical truth table relations.

and similarly,

$$\phi_{\&} = ||\&||:[0,1]\times[0,1] \to [0,1] \tag{9}$$

is a nondecreasing function usually assumed to be bounded above pointwise by the function min, continuous, symm., associative- so that it is unambiguously extendable recursively to any finite number of arguments-and has the boundary truth table values $\phi_{\&}(0,y)=0, \phi_{\&}(1,y)=y$, for all y in [0,1]. An analagous form holds for

$$\phi_{0} = ||\mathrm{or}||:[0,1]\times[0,1] \to [0,1] \ , \tag{10}$$

nondecreasing, etc., bounded below pointwise by max, and having boundary truth table values $\phi_0(0,y)=y, \phi_0(1,y)=1$, for all y in [0,1].
The above functions are called negations (with often the added property of being an involution), t-norms, and t-conorms, respectively. (See [16] for various properties of these operators.)

(ii) If $||\cdot||$ is not truth functional, then the evaluation in eq.(6) does not hold and a more complicated evaluation procedure is valid. One example of this is Probability Logic, where e.g.,

$$||S_1 \text{ or } S_2|| = ||S_1|| + ||S_2|| - ||S_1 \& S_2||, \tag{11}$$

$$||\text{not } S|| = 1-||S|| \ , \tag{12}$$

but in general there is no fixed $\phi_{\&}$, not dependent on any particular S_1 or S_2 such that

$$||S_1 \ \& \ S_2|| = \phi_{\&}(||S_1||,||S_2||),$$

where S_1, S_2, S are any sentences. (See [21] for further discussions concerning truth functional vs. non-truth functional evaluations.)

In addition, for a given set of natural language connectors, more than one semantic evaluation function may be used throughout a given sentence or in certain different sentences.

SOME LANGUAGE OPERATORS AND RELATIONS

In this section some common (but by no means exhaustive) language operators and relations are considered.

(a) Linguistic/logical connectors.
The basic connectors representing negation (not), conjunction (&), disjunction (or), implication (if()then()), have already been introduced. The last could also be defined, as in the classical logic case, in terms of "not" and "or". More compound operators such as "iff" may also be defined. Purely linguistic connectors such as "although" and "but" can be defined entirely in terms of the basic connectors also. For example, "although" may be identified with implication and "but" with conjunction, with some possible modifications.

(b) Hedges.
Hedges are intensifiers or modifiers operating on attributes. If one lets "hedge" represent generically any hedge, such as "extremely", "very", "little", "quite", then any choice of semantic evaluation function $||\cdot||$ leads to the function $\phi_{hedge} = ||hedge||:[0,1] \to [0,1]$. Some controversy exists concerning how to generate spectra of hedges from a neutral hedge, where exponentiation and translation parameter families have been compared empirically as candidates [22],[23],[24]. An alternative, and perhaps more general, approach is to consider first the simple hedges corresponding to integral iterations of conjunction. Thus, for any positive integer j, and any sentence $S=(x\epsilon A)$; x and A are as before :

$$\begin{aligned} S^{(j)} &= \text{"x has property A to the } j^{th} \text{ intensity"} \\ &= \text{"x has property } A^{(j)}\text{"} \\ &= (x\epsilon A)\&(x\epsilon A)\&..\&(x\epsilon A) \quad (j \text{ factors}) \\ &= (x \ \epsilon \ A^{(j)}), \end{aligned} \tag{13}$$

In turn, assuming truth functionality here,

$$||S^{(j)}|| = \phi_{\&}(\phi_A(x),\ldots,\phi_A(x)) \ . \tag{14}$$

However, for the choice $\phi_{\&}$=min, no change in semantic value for the j^{th} intensity is reflected here! On the other hand if $\phi_{\&}$ is an archimedean t-norm such as prod (i.e., ordinary product with respect to its arguments) then it follows from the canonical representation (see e.g. [16], Chapter 2) that there exists a continuous monotone decreasing function $h:[0,1] \to R^+$ with $h(0)\leq +\infty$ and h(1)=0 such that

$$\phi_{\&}(x_1,\ldots x_n) = h^{-1}(\min(h(x_1)+\cdots+h(x_n),h(0))), \tag{15}$$

for all x_i in [0,1], i=1,...,n, n arbitrary positive. (Conversely, any choice of such an h generates an archimedean $\phi_{\&}$ as in (15), where one need only take n=2.) It follows immediately that (15) implies that j in eq.(14) may be replaced by any positive real number, so that (14) becomes

$$||S^{(j)}|| = h^{-1}(\min(j\cdot h(\phi_A(x)),h(0))). \tag{16}$$

Analagous forms may be obtained relative to ϕ_0, and negation as well may be employed. In (16), when $j_0>1$, $S^{(j_0)}$ can be called an intensification, where somewhat arbitrarily, one denotes "very(S)" as $S^{(1)}$, "very very (S)" as "very(very(S))" = $S^{(2)}$, etc. When $j_0<1$, $S^{(j)}$ similarly can be identified with "little of (S)", etc.

Some tie-ins between hedges and quantifiers will be discussed in subsection (e).

(c) Modal operators. (See [25] for background.)
Alethic modality concerns itself with the spectrum- together with negations- of indicativeness. Thus for example: "impossible", "improbable", "possible", "is", "likely", "probable", "certain" is one such collection. Indeed, correspondences have been established between a simple numerical scale of subjective confidences between 0 and 1 and such alethic forms for a number of applications (personal observations). Using negation, the operators of necessity and entailment, among others, may be defined [25]. Deontic modality concerns, analagously, the spectrum of permission or obligation. Other modal families of operators may concern hope, desire, hate, etc.

In any case, a reasonable way to generate such families or spectra of modal operators is to choose some base or anchor within a given family, denoted as modal_0, say, and simply define

$$\text{modal} = \text{hedge}(\text{modal}_0) \tag{17}$$

where hedge is some suitably chosen modifier, as in subsection (e), depending of course on modal. Hence

$$\begin{aligned} S &= \text{modal}(x \ \epsilon \ A) \\ &= (x \ \epsilon \ \text{modal}(A)) \\ &= (x \ \epsilon \ \text{hedge}(\text{modal}_0(A))) \ , \end{aligned} \tag{18}$$

with semantic evaluation, assuming truth functionality,

$$\begin{aligned} ||S|| &= \phi_{modal}(\phi_A(x)) \\ &= \phi_{hedge}(\phi_{modal_0}(\phi_A(x))). \end{aligned} \tag{19}$$

(d) Temporal Operators. (See [26] for the related area of temporal logics.)
Consider first the case for past time operators and in particular the expression

$$S = \text{" y had property A"}. \tag{20}$$

Suppose that A is a generalized subset of domain $X = Y\times R^-$, where Y is some fixed population (an ordinary set) and R^-, the negative reals with zero, represents the flow of time, with the present being identified with t=0. It is also supposed that $\phi_A:X \to [0,1]$ is known.
Thus for any $t \ \epsilon \ R^-$, the sentence

$$S_t = \text{"y had property A at time t"} \tag{21}$$

has the semantic evaluation

$$||S_t|| = \phi_A(y,t) \quad , \tag{22}$$

for any $y \in Y$. Next, identify "was" as a generalized subset of R^-, so that

$$||\text{was}|| = \phi_{was}: R^- \to [0,1] \tag{23}$$

is some monotonically decreasing function with $\phi_{was}(-\infty) = 1$ and $\phi_{was}(0)=0$. At this point it should be remarked that empirical investigations have to be made to determine what the actual membership functions involved in this modeling-and all previously mentioned models-are numerically. Putting together eqs.(21)-(23) yields the reasonable interpretation for eq.(20) :

$$S = \underset{\substack{(\text{over all} \\ t \text{ in } R^-)}}{\text{Or}} \left(S_t \ \& \ (t \in \text{was})\right) \ , \tag{24}$$

which under the usual truth functionality assumptions yields

$$||S|| = \underset{\substack{(\text{over all} \\ t \text{ in } R^-)}}{\phi_0} \left(\phi_\&(\phi_A(y,t), \phi_{was}(t))\right). \tag{25}$$

Note also,that in practise, R^- will be replaced by a suitable discretization, unless ϕ_0=max is chosen. (The problem of extending t-norms and t-conorms to a continuum of arguments is discussed in [16]. A related result may also be found in [12], section 5.) Similar analysis can be carried out for remote past, future, future anterior, and many other temporal relations.

(e) Conditioning and quantification.

Zadeh's contibutions to this area have already been mentioned [5]. See also the discussion of other approaches in [6], pp.138-140. The approach presented here is quite general and reduces to Zadeh's and others for particular evaluations. Let A be a generalized subset of base space X and B a generalized subset of Y. Let quant stand for any quantification involving percentages such as "some","all","few","many","sometimes","often","most","about 3/4","0.456", etc. Let pop be a fixed population of individuals (an ordinary finite set) and suppose that measurement functions $f:\text{pop} \to X$ and $g:\text{pop} \to Y$ are given so that for any z in pop,

$$(z \in A) = \text{"z has attribute A"} = \text{"f(z) has attribute A"}$$
and
$$(z \in B) = \text{"z has attribute B"} = \text{"g(z) has attribute B"}. \tag{26}$$

Furthermore, pop can be considered to be an element-in the ordinary sense- of a superuniversal set *Pop* , the collection of all populations of possible interest. In turn, A and B may also be considered generalized subsets of *Pop* , so that for any member of *Pop*, such as pop, one can define in a reasonable way membership of pop in A and in B as

$$(\text{pop} \in A) = \underset{\substack{(\text{over all} \\ z \in \text{pop})}}{\text{Or}} \left((z \in A) \ \& \ (z \in \text{wt})\right)$$
and
$$(\text{pop} \in B) = \underset{\substack{(\text{over all} \\ z \in \text{pop})}}{\text{Or}} \left((z \in B) \ \& \ (z \in \text{wt})\right), \tag{27}$$

where wt is some generalized subset of pop, representing weighting of importance of each individual for either attribute A or B (assuming here for simplicity that wt is the same for both attributes). If equally likely weighting is desired, $\phi_{wt} \equiv 1/\text{card(pop)}$.Thus, under the usual truth functionality assumptions, it follows that

$$\phi_A(\text{pop}) = ||(\text{pop} \in A)|| = \underset{\substack{(\text{over all} \\ z \in \text{pop})}}{\phi_0} \left(\phi_\&(\phi_A(f(z)), \phi_{wt}(z))\right) \tag{28}$$

with a similar expression holding for $\phi_B(\text{pop})$, where f is replaced by g. Similarly, the evaluation of $\phi_{A \cap B}(\text{pop})$ is given, if no "interaction" is assumed between A and B (see [6] for further details) as:

$$\phi_{A \cap B}(\text{pop}) = ||(\text{pop} \in A \cap B)|| = \underset{\substack{(\text{over all} \\ z \in \text{pop})}}{\phi_0} \left(\phi_\&(\phi_A(f(z)), \phi_B(g(z)), \phi_{wt}(z))\right). \tag{29}$$

The sentences

$$S_1 = \text{"individuals have A, given individuals have B"} = (\text{pop} \in A \mid \text{pop} \in B) \tag{30}$$

and

$$S_2 = \text{"If individuals have B they also have A"} = \text{"If (pop} \in B\text{) then (pop} \in A\text{)"}$$
$$= ((\text{pop},\text{pop}) \in B \times A) = ((\text{pop} \in B) \Rightarrow (\text{pop} \in A)) \tag{31}$$

are slight variations of each other. The first is an example of conditioning, where here conditioning is defined as in [1], Theorem 4. Thus, the semantic evaluation $||S_1||$ satisfies the relation

$$\phi_{A \cap B}(\text{pop}) = ||S_1 \ \& \ (\text{pop} \in B)|| = \phi_\&(||S_1||, \phi_B(\text{pop})). \tag{32}$$

The second is evaluated,as before,as

$$||S_2|| = \phi_\Rightarrow(\phi_B(\text{pop}), \phi_A(\text{pop})), \tag{33}$$

under the usual assumptions.
Then, a sentence such as "Most ships that have long hulls also have maneuvering problems" may be expressed in the general form

$$S_3 = \text{quant}(S_1) \quad \text{or} \quad S_4 = \text{quant}(S_2), \tag{34}$$

leading directly to the evaluations (under truth functionality assumptions)

$$||S_3|| = \phi_{quant}(||S_1||)$$
and
$$||S_4|| = \phi_{quant}(||S_2||), \tag{35}$$

where

$$\phi_{quant} = ||\text{quant}||: [0,1] \to [0,1] \tag{36}$$

is obtained beforehand. For example, ϕ_{Q_1} is conveniently modeled as a unimodal normalized function about 3/4, while ϕ_{Q_2} is a nondecreasing function, being zero over [0,1/2] and then becoming monotone increasing over [1/2,1], where Q_1 = "about 3/4" and Q_2 = "most".

Zadeh's fuzzy cardinality approach to quantification is obtained by choosing $\phi_\&$ = prod, $\phi_{wt}(z) = 1/\text{card(pop)}$, for all z in pop, and by choosing ϕ_0 = bndsum (i.e., for any $v_1, \ldots v_n$ in [0,1], $\text{bndsum}(v_1, \ldots v_n) = \min(1, v_1 + \ldots + v_n)$):

$$||S_3|| = \phi_{quant}\left(\sum_{\substack{(\text{all} \\ z \text{ in pop})}} (\phi_A(f(z)) \cdot \phi_B(g(z))) \Big/ \sum_{\substack{(\text{all} \\ z \text{ in pop})}} (\phi_B(g(z)))\right). \tag{37}$$

Finally, it should be noted that ambiguity arises in the modeling of exact quantifiers. For example, "all" can be approached as above through the function $\phi_{all} = \delta_{\cdot,1}$ (Kronecker delta function for 1) or it can be modeled by the hedge corresponding to the operation $S^{(j)}$ for any sentence S , where here $j \leq \text{card(pop)}$, i.e.,

$$\text{"all z's have A"} = \underset{\substack{(\text{over all} \\ z \text{ in pop})}}{\&} (\phi_A(f(z))) \ . \tag{38}$$

If "softening" is really intended as in"about 5/7" for 5/7" ,"almost all" for "all", "a few" for "there is", etc., then the approach given in this subsection is most appropriate. Conversely, if an exact cardinality is specified as in "at least 2" and is meant literally, then combinatoric considerations have to be made :

"At least two z's in the population which have B , have also A"

$$\text{Or}_{\substack{\text{(over all } z',z'' \\ \text{in pop, } z' \neq z'')}} ((z' \epsilon A) \& (z'' \epsilon A) | (z' \epsilon B) \& (z'' \epsilon B)). \quad (39)$$

(f) Verb and predicative relations.

Three different approaches to the modeling of such relations are presented here.

(i) The relations may be defined operationally-i.e., only directly through a membership function. For example the binary relation "runs to" as in "John runs to the store" can be defined over the domain $X=Y\times Z$, where Y is some relevant human population and Z is a collection of possible objects of the verb "run to".

(ii) The relations may be defined indirectly through the use of measurement functions, as introduced earlier. Thus "gross","fat","small", depending of course on the context, can be directly defined on the domain $R^+\times R^+$ after introducing the natural measurement functions $f:\text{pop} \to R^+$, $g:\text{pop} \to R^+$, representing height in inches and weight in pounds, respectively.

(iii) The relations may be analyzed further, analogous to a dictionary definition of a relatively compound concept in terms of more primative ones. In turn, these relations could be used to form constraints between the components, which would be then modeled. The usefulness of this approach remains to be established.

EXAMPLES ILLUSTRATING SOME OF THE PRINCIPLES

The above stated principles serve as guidelines in the modeling of natural language information. In practice, much ingenuity must be exercised (in a sense, this is an art, based upon intuition) in properly capturing the essence of the meaning of a given sentence. Such will continue to be the case until a universal parsing procedure is discovered (see the comments earlier in the previous subsections)!

Example 1.

Consider the compoundsentence S* in the Introduction:

Let: pop_1 = set of all days of interest , (40)

pop_2 = set of all ocean regions of interest = {V,W,.. } , (41)

pop_3 = set of all submarines of interest , (42)

X=(range of possible temperatures in degrees)
× (range of possible wind velocities in m.p.h.)
× (range of %'s possible representing cloudiness,etc.)
× (range of possible no. representing precip. inten.)
× (range of average maximal visibility in miles)
$\subseteq R\times R^+\times[0,1]\times R^+\times R^+$,representing weather measurements, (43)

Y=(range of wave-chop heights)×(range of max. water vel.) $\subseteq R^+\times R^+$, representing sea state conditions , (44)

with also domains $V,W,.. \subseteq R^2$ (in latitude and longitude). Also define (errorless) measurement functions

$\text{wem}:\text{pop}_1\times\text{pop}_2 \to X$, weather measurement funct. (45)

$\text{ssm}:\text{pop}_1\times\text{pop}_2 \to Y$, sea state meas. function (46)

$\text{loc}:\text{pop}_3 \to V\cup W\cup\cdots$,geolocation meas. function (47)

In particular, for any $z_j \epsilon \text{pop}_j$, j=1,2,

$$\text{wem}(z_1,z_2) = (\text{wem}_1(z_1,z_2),\ldots,\text{wem}_5(z_1,z_2)) \ , \quad (48)$$

so that $\text{wem}_5(z_1,z_2)$ is the av. max. visibility during day z_1 in region z_2 (z_2=V or W).

Next, define generalized set C by, for all $z_j \epsilon \text{pop}_j$, j=1,2 ,

$$\phi_C(z_1,z_2)=\phi_\&(\phi_{\text{poor}}(\text{wem}(z_1,z_2)),\phi_{\text{rel h. turb}}(\text{ssm}(z_1,z_2))) \quad (49)$$

noting that ϕ_{poor} must be modeled and

$$\phi_{\text{rel h. turb}}(x) = \phi_{\text{hedge}_1}(\phi_{\text{normal}}(x)) \ , \text{ all } x\epsilon[0,1] \ , \quad (50)$$

for some properly chosen hedge_1,etc.

Define generlized set D, where for any z_2 in pop_2 ,

$$\phi_D(z_2)=\phi_{\substack{0 \\ \text{(over all} \\ z_3 \text{ in pop}_3)}} \ (||z_3 \text{ was in } z_2||) \ , \quad (51)$$

where

$$||z_3 \text{ was in } z_2||= \phi_{\substack{0 \\ \text{(all t in } R^-)}} (\phi_\&(\phi_{z_2(A)}(\text{loc}(z_3),t),\phi_{\text{was}}(t))). \quad (52)$$

Define generalized set E, for any $z_3',z_3'' \epsilon \text{pop}_3$,

$$\phi_E(z_3',z_3'')=\phi_{\binom{\text{geo}}{\text{match-B}}}(\text{wtd dist}(\text{loc}(z_3'),\text{loc}(z_3''))) \ , \quad (53)$$

where $\phi_{\binom{\text{geo}}{\text{match-B}}}$ arises from ,typically, hypotheses testing of equality of means from gaussian data, and is thus exponential in form (see [16], Chapter 9 for further details relating statistical procedures with this modeling). $\phi_{z_2(A)}(\text{loc}(z_3),t)$, typically may be obtained as the probability function evaluation corresponding to the output of a Kalman filter,for sensor system A.

Next, define generalized set F, where for all $z_1 \epsilon \text{pop}_1$, $z_2 \epsilon \text{pop}_2$, $z_3',z_3'' \epsilon \text{pop}_3$, $\theta \epsilon [0,1]$ -representing possible correlation levels; and for all $\phi_{K_j}:[0,1]\to[0,1]$ for j=1,2,3 ,

$$\phi_F(z_1,z_2,z_3',z_3'',\theta,\phi_{K_1},\phi_{K_2},\phi_{K_3}) =$$

$$\phi_\Rightarrow\{\phi_\&(\phi_C(z_1,z_2),\phi_{K_3}(\phi_{\gtrsim 2}(\text{wem}_5(z_1,z_2)))) \ ,$$
$$(\phi_\Rightarrow\{\phi_\&(\phi_{K_1}(\phi_D(z_2)),\phi_{K_2}(\phi_E(z_3',z_3''))),\phi_{\text{hedge}_2}(\phi_{\text{corr}}(\theta))\}\}) \quad (54)$$

Then define generalized set G, by for all $z_3',z_3'' \epsilon \text{pop}_3$, with z_2=V, and for all ϕ_{K_j} ,j=1,2,3,and all θ,

$$\phi_G(z_3',z_3'',\theta,\phi_{K_1},\phi_{K_2},\phi_{K_3}) =$$

$$\phi_{\substack{\& \\ \text{(all } z_1 \\ \text{in pop}_1)}} (\phi_F(z_1,V,z_3',z_3'',\theta,\phi_{K_1},\phi_{K_2},\phi_{K_3})) \ . \quad (55)$$

Define generalized set H by the conditioning procedure where for all $z_3',z_3'',\theta,\phi_{K_j}$,j=1,2,3, with z_2=W,

$$\phi_H(z_3',z_3'',\theta,\phi_{K_1},\phi_{K_2},\phi_{K_3})=\phi_{\text{most}}\{\text{pop}_1 \epsilon F(\cdot|z_3',z_3'',\theta,\phi_{K_1},\ldots,\phi_{K_3})\}$$

$$=\phi_{\text{most}}(\phi_{\substack{0 \\ \text{(all } z_1 \\ \text{in pop}_1)}} (\phi_\&(\phi_F(z_1,W,z_3',z_3'',\theta,\phi_{K_1},\phi_{K_2},\phi_{K_3}),\phi_{\text{wt}}(z_1)))). \quad (56)$$

Finally, define generalized set L,corresponding to inference rule $S^*(\theta|z_3',z_3'')$, indicating the functional dependencies, as, for all $z_3',z_3'' \epsilon \text{pop}_3$, $\theta \epsilon [0,1]$,

$$\phi_L(\theta,z_3',z_3'') =$$
$$||S^*(\theta|z_3',z_3'')||= \phi_\&(\phi_G(z_3',z_3'',\theta,\phi_{\binom{\text{low}}{\text{effect}}},\phi_{\binom{\text{high}}{\text{effect}}},\phi_{\text{iden}}),$$
$$\phi_H(z_3',z_3'',\theta,\phi_{\binom{\text{im-}}{\text{prove}}}(\phi_{\binom{\text{low-}}{\text{effect}}}(\cdot)),\phi_{\binom{\text{high-}}{\text{effect}}},\phi_N)). \quad (57)$$

with the required models assumed obtainable for the hedges "improve","low effect","high effect",etc.

Example 2.

Consider the compound sentence S** in the Introduction: let all notation be as in Example 1, where required:

Without loss of generality, fix time interval [a,b] compatible with a fixed single day $z_1^* \in pop_1$. Let

pop_4 = set of all surface ships of interest.

Assume that z_4^* is our own ship and z_4 is the target one with $z_4, z_4^* \in pop_4$. Fix also region T^* = region around the Straights of Skagerrak, and let $T^* \in pop_2$. Suppose also that

$$Z=(\text{range of possible hull lengths})\times\cdots \times(\text{range of possible side-insignia colors})\times\cdots \times(\text{range of possible descriptions-locations of prominent objects on ship surface})\times\cdots \subseteq R^+\times\cdots\times(\ldots,\text{red},\text{purple},\ldots)\times\cdots\times(\ldots,(\text{sq.box},\text{front}),\ldots,(\text{oval-dome},\text{rear}),\ldots)\times\cdots , \quad (58)$$

and (errorless) measurement function is given

$$des: pop_4 \to Z \text{ , a description function,} \quad (59)$$

where, for any z_4 in pop_4,

$$des(z_4)=(des_1(z_4),\ldots,des_3(z_4),\ldots,des_{17}(z_4),\ldots). \quad (60)$$

Let "foggy" be a generalized subset of the range of the average maximal visibility in miles, for simplicity. Define generalized set A, where for any times t',t",t"', and z_4,

$$\phi_A(t',t'',t''',z_4)=\phi_\&\Big(\phi_{z_4^* \text{ holds contact over } [t',t'']}(z_4),\ \phi_{\approx 2}(t''-t'),\phi_{small}(t'''-t''),\ \phi_{\text{poster-ior geo}}\big((loc(z_4^*),t''') \,\big|\, wem_5(z_1^*,T^*)\in \text{foggy}\big),\ \phi_{was}(t',t'',t''')\Big), \quad (61)$$

and in turn, define generalized set B, where for all z_4,

$$\phi_B(z_4) = \phi_{\substack{C \\ (\text{over all } t',t'',t''' \text{, with} \\ a\le t'\le t''\le t'''\le b)}}\big(\phi_A(t',t'',t''',z_4)\big). \quad (62)$$

Next, define the generalized set C, where for all z_4,

$$\phi_C(z_4)=\phi_\&\Big\{\phi_{\text{posterior descrip}_3}\big((des_3(z_4)) \,\big|\, (wem_5(z_1^*,T^*)\in\text{foggy})\ \&\ (\text{maybe}(\text{observ}(des_3(z_4))=\text{purple})\big),$$
$$\phi_{\text{posterior descrip}_{17}}\big((des_{17}(z_4)) \,\big|\, (wem_5(z_1^*,T^*)\in\text{ foggy})\ \&\ (\text{maybe}(\text{observ}(des_{17}(z_4))=(\text{oval-dome},\text{rear}))\big),$$
$$\phi_{\text{conditional set of ship names}}\big(des_2(z_4) \,\big|\, des_3(z_4)\ ,\ des_{17}(z_4)\big)\Big\} \quad (63)$$

where des_2 is the naming description such as "Jones", "S.S. Jackson", etc, and where all conditional or conditional posterior generalized sets as above must be appropriately modeled. Then, finally, define the generalized set M, corresponding to information $S^{**}(z_4)$, for all z_4 in pop_4, as

$$\phi_M(z_4) = ||S^{**}(z_4)|| = \phi_\&(\phi_B(z_4),\phi_C(z_4)). \quad (64)$$

Modeling of inference rules such as given in Example 1 and error distribution information as given in Example 2 can be used to extend the usefulness of the PACT or more generally any related combination of evidence procedure which is essentially the semantic evaluation of the disjunction over all nuisance parameter values of the conjunction of all relevant information- in the PACT case, being the conjunction of all relevant inference rules connecting matching levels for attributes with correlation levels and error tables in the form of possibility or membership functions for the attributes in posterior forms, given observed data [1].

SUMMARY AND CONCLUSIONS

An outline has been presented for the modeling and semantic evaluation of linguistic information. The implementation of this depends heavily upon the appropriate modeling of the relevant component membership functions of the generalized sets involved. (See [6], pp. 255-264 for approaches to the latter problem.) Much work remains to be done in the general area.

ACKNOWLEDGMENTS

The author wishes to thank the Naval Ocean System Center IR/IED (Independent Research and Independent Exploratory Development) Office, for support of this and past related work. The author also wishes to express his appreciation to Prof. H.T. Nguyen, New Mexico State University, Las Cruces, for many valuable discussions.

REFERENCES

1. I.R. Goodman (1982), "An approach to the data association and tracking problem through possibility theory", Proc.5 MIT/ONR Wrkshp C^3 Sys., 209-215.
2. I.R. Goodman (1983), "A unified approach to modeling and combining of evidence through random set theory", Proc.6 MIT/ONR Wrkshp C^3 Sys., 42-47.
3. I.R. Goodman (1981), "Applications of a combined probabilistic and fuzzy set technique to the attribute problem in ocean surveillance", Proc. 20 IEEE Conf. Decis. & Cntrl., 1409-1411.
4. L.A.Zadeh(1971), "Quantitative fuzzy semantics", Info. Sci.3, 159-176.
5. L.A. Zadeh (1981), "Test-score semantics for natural languages and meaning representation via PRUF", in Empirical Semantics(B.Rieger,ed.),Bochum Pr.,281-349.
6. D.Dubois & H.Prade(1980), Fuzzy Sets & Systems, Academic Press.
7. E.G.Manes(1982), "A class of fuzzy theories", J.Math. Anal.& Applic. 85, 409-451.
8. I.R.Goodman(1984), "Some new results concerning random sets and fuzzy sets", to be published in Info. Sci.
9. I.R.Goodman(1984), "Identification of fuzzy sets with random sets", to appear in Encyclopedia of Systems & Control (M.Singh,ed.), Pergamon Press.
10. H.L.Wiener,W.W.Willman,I.R.Goodman,J.H.Fullback(1979) Naval Ocean-Surveillance Correlation Handbook,1978, NRL Report 8340, Oct. 31, 1979, Wash.,D.C.
11. I.R.Goodman,H.L.Wiener,W.W.*Willman*(1980),Naval Ocean Surveillance Correlation Handbook,1979, NRL Report 8402,Sept.17, 1980. See also NRL Rpt.8417,Jul.27,1983.
12. I.R.Goodman(1984), "An approach to the target data association problem using subjective and statistical information", Proc. 1984 Amer.Cntrl.Conf.,587-592.
13. R.A.Dillard(1981),Integration of Narrative Processing Data Fusion and Database Updating Techniques in an Automated System,NOSC Tech.480,Oct.29,1981,SanDiego.
14. R.A.Dillard(1983),Representation of Tactical Knowledge Shared by Expert Systems,NOSC.Tech.632,Oct1,1983.
15. L.A.Zadeh(1983),The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems,Memo UCB/ERL M83/41, July 11,1983,El.Res.Lab.Univ.Cal.Berkeley.
16. I.R.Goodman & H.T.Nguyen. Uncertainty Models for Knowledge-Based Systems,North-Holland Pr., to app. 1985.
17. J. Lyons(1979),Semantics,Vols 1 & 2,Cambr.Univ. Press.
18. W.O.Dingwall,ed.(1978),A Survey of Linguistic Science, Greylock Publishers.
19. G.Leech(1974),Semantics,Pelican Books,Great Britain.
20. T.Winograd(1983), Language as a Cognative Process, Vol. 1. Addison-Wesley Co.,Reading, Mass.
21. N.Rescher(1969),Many-Valued Logic,McGraw-Hill Co.
22. G.Lakoff(1973), "Hedges: a study in meaning criteria and the logic of fuzzy concepts", J.Phil.Logic 2,458-508.
23. V. B. Kuz'min (1981), "A parametric approach to description of linguistic values of variables and hedges", Fuzzy Sets & Sys. 6, 27-41.
24. P.J. MacVicar-Whelan (1978), "Fuzzy sets, the concept of height and the hedge VERY", IEEE Trans.Sys.Man & Cybern, vol. SMC-8, no. 6, 507-511.
25. D.P.Snyder(1971), Modal Logic and Its Applications, Van Nostrand Co.
26. N.Rescher & A.Urquhart (1971), Temporal Logic, Springer-Verlag Co., New York.
27. R.Maydole(1973),Many-Valued Logic as a Basis for Set Theory, Ph.D. Dissertation,Boston Univ.,Univ.Mich.Micr.